

**Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Computer Science**



Developing An Intelligent Model To Assist Parsing Holy Quran Text

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Doctor of Philosophy in
Computer Science**

**By
Haval Hajee Ameen**

**Supervised by
Asst. Prof. Dr. AbdulSattar Mohammed Khidhir**

2024 A.D.

1446 A.H.

Abstract

The Holy Quran is of immense importance to many societies due to the position that this book holds for Muslims around the world and the religious teachings it contains. The Holy Quran employs a high-standard Arabic language, which requires analysis and simplification of expressions to enhance comprehension and application of its teachings the process of digitizing and integrating the Holy Quran with computational procedures has facilitated the exploration of the large quantity of information contained in its verses. Through the utilization of these applications, scholars and knowledgeable researchers have effectively established standards that regulate the study of Qur'anic knowledge, therefore broadening the representation of the Quran. In addition, despite possessing the most advanced grammar, organization, and verb conjugation among all-natural languages, Arabic has not gotten much attention. As a result, there is a pressing need for research in this field to help identify comprehensive lexical knowledge. This study focuses on the analysis of sentences found in the holy Quran. After dividing all verses or sentences into grammatical sections that represent a letter or word, or the sentence may be complete in one word and contain more than one grammatical section. A neural network was employed. The token word attribute is trained using a neural network input word characteristic, which consists of two layers. This study utilized a knowledge base of 160 elements, from which 26 features specifically relevant to the grammatical case were chosen. The first consisted of an input dataset with 128,211 rows, while the second consisted of a dataset with 99,540 rows. Several separate trials were conducted in each group, which were designated as Case I, Case II, and Case III. Which depends on the number of grammatical sections entered, which are 4,7 and 10. each group has 80% for training and 20% for testing. The results are as follows the neural network's performance might vary depending on the number of nodes utilized., the best results for accuracy were and it was 99.02%.



وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم علوم الحاسوب

تطوير نموذج ذكي للمساعدة في اعراب نصوص من القرآن الكريم

اطروحة مقدمة

الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة دكتوراه فلسفة في
علوم الحاسوب
من قبل

هفال حجي امين محمود

بإشراف
ا.م.د. عبدالستار محمد خضر

الخلاصة

إن القرآن الكريم له أهمية كبيرة لدى العديد من المجتمعات وذلك لمكانته لدى المسلمين في مختلف أنحاء العالم وما يحتويه من تعاليم دينية. إن القرآن الكريم يستخدم لغة عربية عالية المستوى، الأمر الذي يتطلب تحليل وتبسيط التعابير لتعزيز فهم وتطبيق تعاليمه، وقد سهلت عملية رقمنة القرآن الكريم ودمجه بالإجراءات الحسابية استكشاف الكم الهائل من المعلومات الواردة في آياته. ومن خلال الاستفادة من هذه التطبيقات، نجح العلماء والباحثون المطلعون في إرساء معايير فعالة تنظم دراسة المعرفة القرآنية، وبالتالي توسيع تمثيل القرآن.

بالإضافة إلى ذلك، وعلى الرغم من امتلاك اللغة العربية لأحدث قواعد النحو والتنظيم وتصريف الأفعال بين جميع اللغات الطبيعية، إلا أنها لم تحظ باهتمام كبير. ونتيجة لذلك، هناك حاجة ملحة للبحث في هذا المجال للمساعدة في تحديد المعرفة المعجمية الشاملة. تركز هذه الدراسة على تحليل الجمل الموجودة في القرآن الكريم.

وبعد تقسيم جميع الآيات أو الجمل إلى مقاطع نحوية تمثل حرفاً أو كلمة، أو قد تكون الجملة كاملة بكلمة واحدة وتحتوي على أكثر من مقطع نحوي. وتم استخدام شبكة عصبية. ويتم تدريب سمة الكلمة الرمزية باستخدام خاصية إدخال الكلمة في الشبكة العصبية، والتي تتكون من طبقتين.

وقد استفادت هذه الدراسة من قاعدة معرفية مكونة من 160 عنصراً، تم اختيار 26 سمة منها ذات صلة خاصة بالحالة النحوية. وتألقت الأولى من مجموعة بيانات إدخال تحتوي على 128211 سطرًا، بينما تتكون الثانية من مجموعة بيانات تحتوي على 99540 سطرًا. وتم إجراء عدة تجارب منفصلة في كل مجموعة، والتي تم تحديدها على أنها الحالة الأولى والحالة الثانية والحالة الثالثة. والتي تعتمد على عدد المقاطع النحوية المدخلة، وهي 4 و 7 و 10. ولكل مجموعة 80% للتدريب و 20% للاختبار. وكانت النتائج على النحو التالي: قد يختلف أداء الشبكة العصبية اعتمادًا على عدد العقد المستخدمة، وكانت أفضل النتائج من حيث الدقة 99.02%.